rev. version of 06/25/2004

Learning to Network

Brian Skyrms
Robin Pemantle

1. Introduction

In species capable of learning, including our own, individuals can modify their behavior by some adaptive process. Important classes of behavior - mating, predation, coalitions, trade, signaling, and division of labor - involve interactions between individuals. The agents involved learn two things: with *whom to interact* and *how to act*. That is to say that adaptive dynamics operates both on structure and strategy.

In an interaction individuals actualize some behavior, the behavior of the individuals jointly determines the outcome of the interaction, and the consequences for the individuals motivate learning. At this high level of abstraction, we can model interactions as games. The relevant behaviors of individuals are called strategies of the game, and the strategies of the players jointly determine their payoffs. Payoffs drive the learning dynamics. (Skyrms and Pemantle, 2000).

If we fix the interaction structure in this abstract scheme, we get models of the evolution of strategies in games played on a fixed structure. An interaction structure need

not be deterministic. In general, it can be thought of as a specification of the probabilities of interaction with other individuals. By far the most frequently studied interaction structure is one in which the group of individuals is large and individuals interact at random. That is to say that each individual has equal probability of interacting with every other individual in the population. Among a list of virtues of this model, mathematical tractability must come near the top. At another end of the spectrum we have models where individuals interact with their neighbors on a torus, or a circle, or (less frequently) some other graphical structure. [Ellison (1993) , Nowak and May (1992), Hegselmann (1996), Alexander (2000)] . Except in the simplest cases, these models sacrifice mathematical tractability to gain realism, and computer simulations have played an important role in their investigation. These two extreme models, however, can have quite different implications for the evolution of behavior. In large, random encounter settings cooperators are quickly eliminated in interactions with a Prisoner's Dilemma structure. Comparable local interaction models allow cooperators to persist in the population.

If we fix the strategies of individuals and let the interaction structure evolve we get a model of interaction network formation. Evolution of structure is less well-studied than evolution of strategies, and is the main focus of this paper. Most current research on theory of network formation takes the point of view that networks are modeled as graphs or directed graphs, and network dynamics consists of making and breaking of links. [Jackson and Watts (2002), Bala and Goyal, (2000)] In taking an interaction structure to be a specification of probabilities of interaction rather than a graphical structure, we take a more general view than most of the literature (but see Kirman (1997) for a point of view

close to that taken here.) It is possible that learning dynamics may drive these probabilities to zero or one and that a deterministic graphical interaction structure may crystallize out, but this will be treated as a special case. We believe that this probabilistic approach can give a more faithful account of both human and non-human interactions. It also makes available a set of mathematical tools that do not fit the coarser picture of making or breaking deterministic links in a graphical structure.

Ultimate interest resides in the general case where structure and strategy co-evolve. These may be modified by the same or different kinds of learning. They may proceed at the same rate or different rates. The case where structure dynamics is slow and strategy dynamics is fast may approximate more familiar models where strategies evolve on a fixed interaction structure. The opposite case may be close to that of individuals with fixed strategies (or phenotypes) learning to network. In between, there is a very rich territory waiting to be explored. We will close this paper with a discussion of the co-evolution of structure and strategy a game which one of us has argued is the best simple prototype of the problem of instituting a social contract.[Skyrms, 2004]. Whether coevolution of structure and strategy supports or reverses the conventional wisdom about equilibrium selection in this game, depends on the nature and relative rates of the two learning processes.

2. Learning

Learning can be divided into two broad categories: (1) belief learning in which the organism forms belief or internal representations of the world and uses these to make decisions, and (2) reinforcement learning, where the organism increases the probability of acts that have been rewarded and decreases the probability of those that have not been rewarded. Ultimately the distinction may not be so clear cut, but it is useful for a categorization of learning theories. In the simplest belief learning model, Cournot dynamics, an individual assumes that others will do what they did last time and performs the act that has the highest payoff on that assumption. More sophisticated individuals might form their beliefs more carefully, by applying inductive reasoning to some or all of the available evidence. Less confident individuals might hedge their bet on Cournot dynamics with some probabilistic version of the rule. Strategically minded individuals might predict the effect of their current choice on future choices of the other agents involved, and factor this into their decision. Humans, having a very large brain, can do all of these things but often they do not bother [Suppes and Atkinson (1960), Roth and Erev (1995), Erev and Roth (1998), Busemeyer and Stout (2002), Yechiam, and Busemeyer (preprint).]

Reinforcement learning does not require a lot of effort, or a large brain, or any brain at all. In this paper we will concentrate on reinforcement learning, although we will also touch on other forms of learning. Specifically, we apply a mathematical model in which the probability of an act is proportional to the accumulated rewards from performing that act. [Herrnstein (1970), Roth and Erev (1995)]. Following Luce (1959), the learning model can be decomposed into two parts, a reinforcement dynamics, in which

weights or propensities for acts evolve, and a response rule, which translates these weights into probabilities of acts. If we let weights evolve by adding the payoff gotten to the weight of the act chosen, and let our probabilities be equal to the normalized weights (Luce's linear response rule), we get the basic Herrnstein-Roth-Erev dynamics.

There are alternative models of reinforcement learning that could be investigated in this setting. In a path-breaking study, Suppes and Atkinson (1960) applied stimulus sampling dynamics to learning in two-person games. Borgers and Sarin (1997) have investigated dynamics of Bush and Mosteller (1955) dynamics in a game-theoretic setting. Instead of the Luce's linear response rule of normalizing the weights, some models use a logistic response rule. Bonacich and Liggett (2004) apply Bush-Mosteller learning in a setting closely resembling our own. They get limiting results that are closely connected to this of the discounted model of Friends II in Skyrms and Pemantle (2000). Liggett and Rolles (in press) generalize the results of Bonacich and Liggett to an infinite space of agents. We, however, will concentrate attention on the basic Herrnstein-Roth-Erev dynamics and on a "slight" variation on it.

Erev and Roth (1997) suggest modifying the basic model by discounting the past to take account of "forgetting". At each time period, accumulated weights are multiplied by some positive discount factor less than one, while new reinforcements are added at full strength. Discounting is a robust phenomenon in experimental studies of reinforcement learning, but there seems to be a great deal of individual variability with reported discount factors ranging from .5 to .99. [Erev and Roth (1997), Busemeyer and Stout (2002),

Yechiam and Busemeyer (2004), Goeree and Holt (forthcoming)]. Discounting changes the limiting properties of the learning process radically. We will see that within the reported range of individual variability, small variations in the discount rate can lead to large differences in predicted observable outcomes in interactive learning situations.

3. Two-Person Games with Basic Reinforcement Learning

We begin by investigating basic (undiscounted) reinforcement learning in simple two-person interactions. The following model was introduced in Skyrms and Pemantle (2000). Each day each individual in a small group wakes up and decides to visit someone. She decides by chance, with the chance of visiting anyone else in the group being given by normalized weights for that individual. (We can imagine the process starting with some initial weights; they can all be set to one to start the process with random encounters.) The person selected always accepts, and there is always time enough in the day for all selected interactions to take place. In a group of ten, if Jane decides to visit someone and the other nine all happen to decide to visit Jane, she has a total of ten interactions in that day. Each interaction produces a payoff. At the end of the day, each individual updates her weights for every other individual by adding the payoffs gotten that day from interactions with that individual. (Obvious variations on the basic model suggest themselves, but we confine ourselves here to just this model applied to different kinds of interactions.) Initially, we investigate baseline cases where individuals have only the choice of with whom to interact, and interactions always produce payoffs in the same way. Then we build on the results for

these cases to analyze interactions in the stag hunt game, in which different agents can have different acts and the combination of acts determines the payoffs.

Consider two games of "Making Friends." In Friends I the visitor is always treated well, and gains a payoff of 1, while the host goes to some trouble but also enjoys the encounter, for a net payoff of zero. In Friends II the visitor and host are both equally reinforced, with a payoff of 1 going to each. We start each learning process with each individual having initial weights of one for each other individual, so that our group begins by interacting at random. It is easy to run computer simulations of the Friends I and Friends II processes, and it is a striking feature of such simulations that in both cases non-random interaction structure rapidly emerges. Furthermore, rerunning the processes from the same starting point seems to generate different structure each time. In this setting, we should expect the emergence of structure without an organizer, or even an explanation in terms of payoff differences. The state of uniform random encounters with which we started the system does not persist, and so must count as a very artificial state. Its use as the fixed interaction structure in many game theoretic models is therefore extremely suspect.

We can understand the behavior of the Friends I process if we notice that each individual's learning process is equivalent to a Polya urn. We can think of him as having an urn with balls of different colors, one color for each other individual. Initially there is one ball of each color. A ball is chosen (and returned), the designated individual is visited. Because visitors are always reinforced, another ball of the same color is added to the urn.

Because only visitors are reinforced, balls are not added to the urn in any other way. (Philosophers of science will be familiar with the Polya urn because of its equivalence with Bayes-Laplace inductive inference.) The Polya urn converges to a limit with probability one, but it is a random limit with uniform distribution over possible final probabilities. Anything can happen, and nothing is favored! In Friends I the random limit is uniform for each player, and makes the players independent. [Skyrms and Pemantle 2000, Theorem 1]. All interaction structures are possible in the limit, and the probability that the group converges to random encounters is zero.

In Friends II, both visitor and host are reinforced and so the urns interact. If someone visits you, you are reinforced to visit him - or to put it graphically, someone can walk up to your door and put a ball of his color in your urn. This complicates the analysis. Nevertheless, the final picture is quite similar. The limiting probabilities must be *symmetric*, that is to say X visits Y with the same probability that Y visits X, but subject to this constraint and its consequences anything can happen. [Skyrms and Pemantle (2000) Theorem 2].

So far, the theory has explained the surprising results of the simulations, but a rather special case of Friends II provides a cautionary contrast. Suppose that there are only three individuals. (What we are about to describe is much less likely to happen if the number of individuals is a little larger.) Then the only way we can have symmetric visiting probabilities is if each individual visits the other two each with probability one-half. Then the previous theorem implies that in this case the process must converge to these

probabilities. In simulations this sometimes happens rapidly. However, there are other trials in which the system appears to be converging to a state in which individual A visits B and C equally, but B and C always visit A and never each other. You can think of individual A as "Ms. Popular." The system was observed to stay near such a state for a long time (5,000,000 iterations of the process.)

This apparent contradiction is resolved in Pemantle and Skyrms (2005), using the theory of stochastic approximation. For the basic Herrnstein-Roth-Erev model, there is an underlying deterministic dynamics that can be obtained from the expected increments of the stochastic process. This deterministic dynamics has four equilibria - one in which each individual visits the others with equal probability and the other three having A, B, and C respectively as "Ms. Popular." The symmetric equilibrium is strongly stable - an attractor - while the "Ms. Popular" equilibria are unstable saddle points. The system must converge to the symmetric equilibrium. It cannot converge to one of the unstable saddles, but if in the initial stages of learning it wanders near a saddle it may take a long time to escape because the vector pushing it away is very small. This is what happens in the anomalous simulations. There is a methodological moral here that we will revisit in the next section. Simulations may not be a reliable guide to limiting behavior and limiting behavior is not necessarily all that is of interest.

The Making Friends games provide building blocks for analyzing learning dynamics where the interactions are games with non-trivial strategies. Consider the two-person Stag Hunt. Individuals are either Stag Hunters or Hare Hunters. If a Stag Hunter interacts with

a Hare Hunter no Stag is caught and the Stag Hunter gets zero payoff. If a Stag Hunter interacts with another Stag Hunter the Stag is likely caught and the hunters each get a payoff of one. Hare Hunting requires no cooperation, and its practitioners get a payoff of .75 in any case. The game is of special interest for social theory because cooperation is both mutually beneficial and an equilibrium, but it is risky [Skyrms (2004)]. In game theoretic terminology, Stag hunting is payoff dominant and Hare hunting is risk dominant. In a large population composed of half Stag Hunters and half Hare Hunters with random interactions between individuals, the Hare Hunters would get an average payoff of .75 while the Stag Hunters would only get an average payoff of .50. The conventional wisdom is that in the long run evolution will strongly favor Hare hunting, but we say that one should consider the possibility that the players *learn to network*.

We use exactly the same model as before, except that the payoffs are now determined by the individuals' types or strategies: Hunt Stag or Hunt Hare. We start with an even number of Stag Hunters and Hare Hunters. Theory predicts that, in the limit, Stag Hunters always visit Stag Hunters and Hare Hunters always visit Hare Hunters [Skyrms and Pemantle (2000) Th. 6]. Simulation confirms that such a state is approached rapidly. Although on rational choice grounds Hare Hunters "should not care" whom they visit, they cease to be reinforced by visits from Stag Hunters after Stag Hunters learn not to visit them. Hare Hunters continue to be visited by other Hare Hunters, so all the differential learning for Hare Hunters takes place when they are hosts rather than visitors. Once learning has sorted out Stag Hunters and Hare Hunters so that each group only

interacts with its own members, each is playing Friends II with itself and previous results characterize within-group interaction structure.

Now Stag Hunters prosper. Was it implausible to think that Stag Hunters might find a way to get together? If they were sophisticated, well-informed, optimizing agents they would have gotten together right away! Our point is that it doesn't take much for Stag Hunters to get together. A little bit of reinforcement learning is enough.

4. Clique Formation with Discounting the Past

Adding a little discounting of the past is a natural and seemingly modest modification of the reinforcement process. However, it drastically alters the limiting behavior of learning. If the Polya urn, which we used in the analysis of Friends I, is modified by discounting the past the limiting result is that after some time (can't say when) there will be one color (can't say which) that will always be picked. Discounting the past, no matter how little the discounting, leads to deterministic outcomes. This is also true when we learn to network. Discounting the past leads to the formation of cliques, whose members never interact with members of alternative cliques. Why then, did we even bother to study learning without discounting? We will see that
if discounting is small enough, learning with discounting may, for long periods of time, behave like learning without discounting.

The effects of adding discounting to the learning process are already apparent in two-person interactions [Skyrms and Pemantle (2000)], but they are more interesting in multi-person interactions. Here we discuss two three-person interactions, Three's Company (a uniform reinforcement counterpart to Friends II), and a Three-Person version of the Stag Hunt. Every day, each individual picks two other individuals to visit to have a three-person interaction. The probability of picking a pair of individuals is taken to be proportional to the product of their weights. The payoff that an individual receives from a three-person interaction is added to her weights for each of the other two participants. We again start the learning process with random interaction. Everyone begins having weight one for everyone else.

In Three's Company, as in Friends II, everyone is always reinforced in every interaction. Everyone gets a payoff of one. No matter what the discount rate, the limiting result of discounted learning is clique formation. For a population size of six or more, the population will break up into cliques of size 3, 4, or 5. Each member of a given clique chooses each other member of that clique with positive limiting relative frequency. For each member of a clique, there is a finite time after which she does not choose outsiders. All such cliques - that is each partition of the population into sets of size 3, 4, and 5, has positive probability of occurring. [Pemantle and Skyrms (in press) Th. 4.1.]

Simulations at a discount rate of .5 conform to theory. A population of 6 always broke into two cliques of size 3, with no interactions between cliques. As we discount less - keeping more of the past weights - we see a rapid shift in results. Multiplying past

weights by .6, led to formation of 2 cliques in 994/1000 trials; by .7 in 13/1000; by .8 in none. (We ran the process for 1,000,000 time steps and rounded interaction probabilities to two decimal places.) Writing the discount factor by which past payoffs are multiplied as (1-x), we can say that simulation says that clique formation occurs reliably for large x, but not at all for small x with a large transition taking place between x=.4 and x=.3. The theory says that clique formation occurs for any positive x.

This apparent conflict between theory and simulation is resolved in Pemantle and Skyrms (2004), where it is shown that time to clique formation increases exponentially in 1/x as the discount factor (1-x) approaches 1. The behavior of the process for observable finite sequences of iterations is highly sensitive to the discount parameter, within ranges that fall within the individual variability that has been reported in the experimental literature. When x is close to 1, discounted reinforcement learning behaves for long periods of time like undiscounted learning in which clique formation almost never occurs.

Three's Company, like Friends II, is important because it arises naturally in the analysis of less trivial interactions. Consider a Three-Player Stag Hunt [Pemantle and Skyrms (in press)]. Pairs of individuals are chosen, and weights evolve, just as in Three's Company, but the payoffs depend on the types of players. If three Stag Hunters interact, they all get a payoff of 4, but a Stag Hunter who has at least one Hare Hunter in his trio gets nothing. (In a random encounter setting, Stag Hunting is here even more risky that in the two person case.) Hare hunters always get a payoff of 3.

In the limit Stag Hunters learn to always visit other Stag Hunters but, unlike some other limiting results we have discussed, this one is attained very rapidly. With 6 Stag Hunters and 6 Hare Hunters and a discount rate of .5, the probability that a stag hunter will visit a hare hunter usually drops below half a percent in 25 interactions. In 50 iterations this always happened in 1000 trials, and this remains true for values of x between .5 and .1. For x=.01, 100 iterations suffices and 200 iterations are enough if x=.001.

Once Stag Hunters learn to visit Stag Hunters, they are essentially playing a game of Three's Company among themselves. They may be visited by Hare Hunters, but these visits produce no reinforcement for the Stag Hunters and so do not alter their weights. Stag Hunters then form cliques of size 3, 4, or 5 among themselves. This will take a long time if the past is only slightly discounted.

There is a tendency for Hare Hunters to learn to visit Hare Hunters after the Stag Hunters learn not to visit them, but because of the discounting it is possible for a Hare Hunter to be frozen in a state of visiting one or two Stag Hunters. This is a real possibility when the past is heavily discounted. At x=.5, at least one Hare Hunter interacted with a Stag Hunter (after 10,000 iterations) in 384 out of 1000 trials. This dropped to 6/1000 for x = .2 and to 0 for x =.1. Hare Hunters who are no trapped into interactions with Stag Hunters eventually end up playing Three's Company among themselves and also form cliques of size 3, 4, and 5.

5. Coevolution of Structure and Strategy

So far we have concentrated on the dynamics of interaction, because we believe that it has not received as much attention as it deserves. The full story involves co-evolution of both interaction structure and strategy. Depending on the application, these may involve the same or different adaptive dynamics and they may evolve at the same or different rates. We will illustrate this with two different treatments of the two-person Stag Hunt.

To the two-person Stag Hunt of section 3, we add a strategy revision process based on imitation. This *reinforcement-imitation* model was discussed in Skyrms and Pemantle (2000). With some specified probability, an individual wakes up, looks around the whole group, and if some strategy is prospering more than his own, switches to it. Individual's probabilities are independent. If imitation is fast relative to structure dynamics, it operates while individuals interact more or less at random and Hare Hunters will take over more often than not. If imitation is slow, stag hunters find each other and prosper, and then imitation slowly converts Hare Hunters to Stag Hunters (who quickly learn to interact with other Stag Hunters).

Simulations show that in intermediate cases, timing can make all the difference. We start with structure weights equal to 1 and vary the relative rates of the dynamics by varying the imitation probability. With "fast" imitation (pr = .1) 78% of the trials ended up with everyone converted to Hare Hunting and 22% ended up with everyone converted to Stag Hunting. Slower imitation (pr = .01) almost reversed the numbers, with 71% of the trials ending up All Stag and 29% ending up All Hare. Fluid network structure coupled with slow strategy revision reverses the orthodox prediction that Hare Hunting (the risk dominant equilibrium) will take over.

(This conclusion remains unaffected if we add discounting to the learning dynamics for interaction structure. Discounting the past simply means that Stag Hunters find each other more rapidly. No matter how Hare Hunters end up, Stag Hunters are more prosperous. Imitation converts Hare Hunters to Stag Hunters.)

The foregoing model illustrates the combined action of two different dynamics, reinforcement learning for interaction structure and imitation for strategy revision. What happens if both processes are driven by reinforcement learning? In particular, we would like to know whether the relative rates of structure and strategy dynamics still make the same difference between Stag Hunting and Hare Hunting. In this *Double Reinforcement* model, each individual has two weight vectors, one for interaction propensities and one for propensities to either Hunt Stag or Hunt Hare. Probabilities for whom to visit and what to do are both gotten by normalizing the appropriate weights. Weights are updated by adding the payoff from an interaction to both the weight for the individual involved and

to the weight for the action taken. Relative rates of the two learning processes can be manipulated by changing the magnitude of the initial weights.

In the previous models we started the population off with some Stag Hunters and some Hare Hunters. That point of view is no longer correct. The only way one could be deterministically a Stag Hunter would be if he started out with zero weight for Hare Hunting, and then he could never learn to hunt Stag. We have to start out individuals with varying propensities to hunt Hare and Stag. There are various interesting choices that might be made here; we will report some simulation results for one. We start the a group of 10, with 2 confirmed Stag Hunters (weight 100 for Stag, 1 for Hare), 2 confirmed Hare Hunters (weight 100 for Hare, 1 for Stag), and 6 undecided guys (weights 1 for Stag and 1 for Hare. Initial weights for interaction structure were all equal, but their magnitude was varied from .001 to 10, in order to vary the relative rates of learning structure and strategy. The percent of 10,000 trials that ended up All Stag or All Hare (after 1,000,000 iterations) for these various settings are shown in figure 1. As before, fluid interaction structure and slow strategy adaptation favor Stag Hunting, while the reverse combination favors Hare Hunting.

(figure 1 here)

In both reinforcement-imitation and double reinforcement models of the coevolution of structure and strategy a fluid network structure shifts the balance form the risk dominant Hare Hunting equilibrium to the cooperative Stag Hunt.

6. Why Dynamics?

Classical, pre-dynamic, game theory would approach the problem differently. The whole group of 10 individuals is playing a 10-person game. A move consists in choosing both a person to play with and a strategy. We can just identify the Nash equilibria of this large game. None are strict. The pure equilibria fall into two classes. One class has everyone hunting Stag and every possible interaction structure. The other has everyone hunting Hare and every possible interaction possible interaction structure. (There are also mixed equilibria with every possible interaction structure.) From this point of view, interaction structure does not seem very important. If you ignore dynamics you miss a lot.

References:

Alexander, J. M. (2000) Evolutionary Explanations of Distributive Justice. *Philosophy of Science* 67: 490-516.

Bala, V. and Goyal, S. (2000) A Non-Cooperative Model of Network Formation. *Econometrica* 68, 1181-1229.

Bonacich, P. and Liggett, T. (2003) Asymptotics of a Matrix-Valued Markov Chain Arising in Sociology. *Stochastic Processes and Their Applications* 104, 155-171.

Busemeyer, J. and Stout, J. (2002) A Contribution of Cognitive Decision Models to Clinical Assessment: Decomposing Performance on the Bechara Gambling Task. *Psychological Assessment* 14, 253-262.

Bush, R. and Mosteller, F. (1955) *Stochastic Models of Learning*. John Wiley & Sons, New York.

Ellison, G. (1993) Learning, Local Interaction, and Coordination. *Econometrica* 61,1047-1071.

Erev, I. and Roth, A. (1998) Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique Mixed Strategy Equilibria. *American Economic Review* 88, 848-881.

Goeree, J. K. and Holt, C. A. (forthcoming) Learning in Economics Experiments. In *Encyclopedia of Cognitive Science*. Macmillan: New York.

Hegselmann, R. (1996) Social Dilemmas in Lineland and Flatland. In *Frontiers in Social Dilemmas Research* Ed. W. Liebrand and D. Messick. Berlin: Springer, 337-362.

Herrnstein, R. J. (1970) On the Law of Effect. *Journal of the Experimental Analysis of Behavior*. 13, 243-266.

Jackson, M. and Watts, A. (2002) On the Formation of Interaction Networks in Social Coordination Games. *Games and Economic Behavior* 41: 265-291.

Kirman, A, (1997) The Economy as an Evolving Network. *Journal of Evolutionary Economics* 7: 339-353.

Liggett, T. M. and Rolles, S. (in press) An Infinite Stochastic Model of Social Network Formation. *Stochastic Processes and Their Applications.*

Luce, R. D. (1959) *Individual Choice Behavior*. John Wiley and Sons, New York.

Pemantle, R. and Skyrms, B. (2004) Time to Absorption in Discounted Reinforcement Models. *Stochastic Processes and Their Applications*. 109: 1-12.

Pemantle R. and Skyrms, B. (in press) Network Formation by Reinforcement Learning: The Long and the Medium Run. *Mathematical Behavioral Sciences.*

Pemantle, R. and Skyrms, B. (in preparation) Reinforcement Schemes May Take a Long Time to Exhibit Limiting Behavior.

Roth, A. and Erev, I. (1995) Learning in Extensive Form Games: Experimental Models and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior* 8, 14-212.

Skyrms, B. (2004) *The Stag Hunt and the Evolution of Social Structure*. Cambridge, New York.

Skyrms, B. and R. Pemantle (2000) A Dynamic Model of Social Network Formation. *Proceedings of the National Academy of Sciences of the USA*. 97: 9340-9346.

Suppes, P. and Atkinson, R. (1960) *Markov Learning Models for Multiperson Interactions*. Stanford University Press, Palo Alto.

Yechiam, E. and Busemeyer, J. R. (1994) Comparison of Basic Assumptions Embedded in Learning Models for Experienced Based Decision-Making. *preprint*.

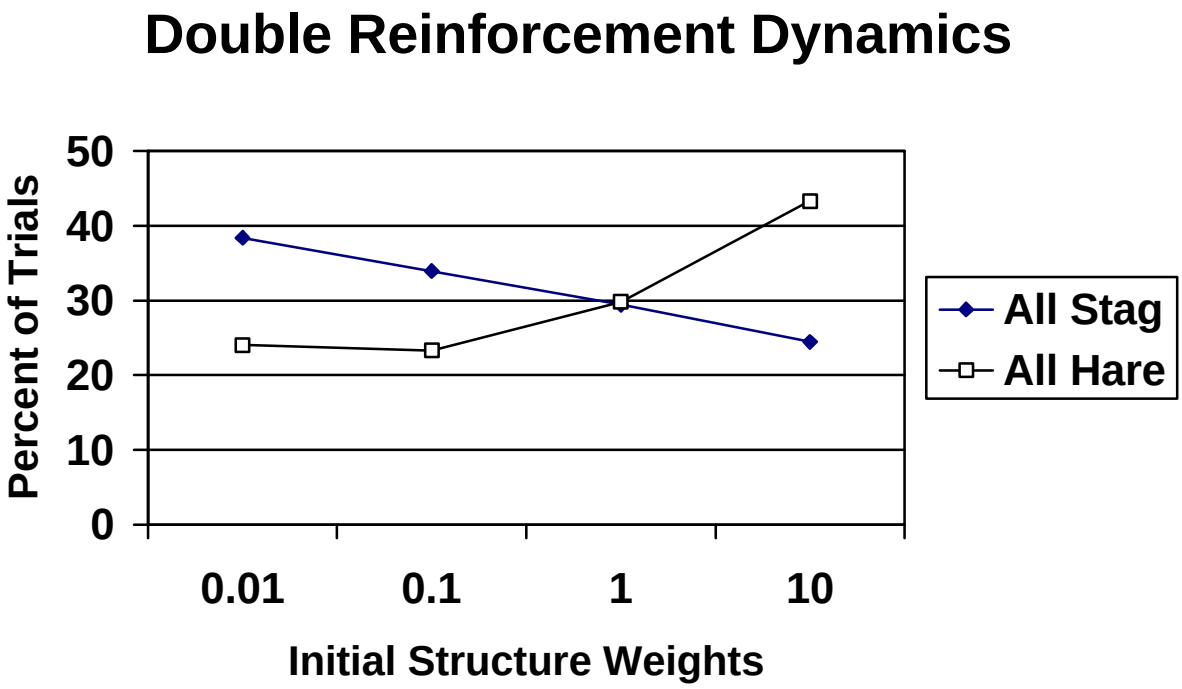

figure 1: Stag Hunt with Reinforcement Dynamics for both Strategy and Structure